Investigation of diagnostic value of artificial intelligence systems in the diagnosis of breast cancer based on histopathological images using Meta-MUMS DTA tool

Massoud Sokouti [1], Ramin Sadeghi [1], Saeid Pashazadeh [2], Saeid Eslami Hasan Abadi [3], Mohsen Sokouti [4], Morteza Ghojazadeh [5], Babak Sokouti [6]**

(1) Nuclear Medicine Research Center, Mashhad University of Medical Sciences, Mashhad, Iran
(2) Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran
(3) Department of Medical Informatics, Faculty of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran
(4) Department of Cardiothoracic Surgery, Tabriz University of Medical Sciences, Tabriz, Iran
(5) Research Center of Evidence-Based Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
(6) Biotechnology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran

CORRESPONDING AUTHOR: *Babak Sokouti, Biotechnology Research Center, Tabriz University of Medical Sciences, Tabriz, Iran. Email: b.sokouti@gmail.com; sokoutib@tbzmed.ac.ir; and Ramin Sadeghi, Nuclear Medicine Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. Email: raminsadeghi1355@yahoo.com*

DOI: ***10-2427/13313***
Accepted on April 27, 2020

ABSTRACT

Background: Various artificial intelligence systems are available for diagnosing breast cancer based on histopathological images. Assessing the performance of existing methodologies for breast cancer diagnosis is vital.
Methods: The SCOPUS database has been searched for studies up to December 15, 2018. We extracted the data, including "true positive," "true negative," "false positive," and "false negative". The pooled sensitivity, pooled specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, area under the curve of summary receiver operating characteristic curve were useful in assessing the diagnostic accuracy. Egger's test, Deeks' funnel plot, SVE (Smoothed Variance regression model based on Egger's test), SVT (Smoothed Variance regression model based on Thompson's method), and trim and fill methodologies were essential tests for publication bias identification.
Results: Three studies with eight approaches from thirty-seven articles were found eligible for further analysis. A sensitivity of 0.95, a specificity of 0.78, a PLR of 7525, an NLR of 0.06, a DOR of 88.15, and an AUC of 0.953 showed high significant heterogeneity; however, the reason was not the threshold effect. The publication bias was detected by SVE, SVT, and trim and fill analysis.
Conclusion: The artificial intelligent (AI) systems play a pivotal role in the diagnosis of breast cancer using histopathological cell images and are important decision-makers for pathologists. The analyses revealed that the overall accuracy of AI systems is promising for breast cancer; however, the pooled specificity is lower than pooled sensitivity. Moreover, the approval of the results awaits conducting randomized clinical trials with sufficient data.

Key words: Meta-analysis, Diagnosis, Breast Cancer, Artificial Intelligence Systems, Cell Images, Histopathology, Accuracy.

INTRODUCTION

According to the report announced by World Health Organization (WHO) on 12 September 2018, one of the most common cancers ranked after lung cancer is breast cancer, impacting almost the same number of patients (2.1 million cases) in 2018 (4). And, the most recent estimate revealed that breast cancer would be the cause of death among women by 627,000 in the current year (4). Despite the gradually increasing trend of diagnosed cases annually, early diagnosis and screening are of "first-must-to-do" strategies. Generally, cancers are the outcome of the transformation of healthy cells to malignant cells developing through pre-cancerous to cancerous multi-stages (5-7). For this purpose, several methodologies and approaches, whether they are using clinical or computational methods, have been taken into the researchers' consideration. Among those, the current study has focused on the machine learning approaches, mathematical models trained and tested/validated on training and validating datasets, for classification of the stages of breast cancer based on the features derived from analyzing the histopathological images.

An artificial intelligence decision system works such as the brain of a clinician or professional medical doctor whose prognosis or diagnosis approach employs the symptoms or properties known for the disease, particularly breast cancer. When the histopathology cell images are obtained from several specific patients, some essential features will be derived using advanced image processing algorithms (8-12). Then, the extracted features will be used for training the artificial intelligence system, that may include globally well-known algorithms such as artificial neural networks (ANNs) (13, 14), support vector machines (SVMs) (15, 16), and random forest (RF) (17, 18). These algorithms may then be optimized further using genetic algorithm (GA) (19-22), ant colony optimization (ACO) (23-25), and particle swarm optimization (PSO) (26-30).

Although there are many types of research carried out in this field; however, due to the lack of publicly available developed methodologies and their datasets, no critical assessments can be done regarding their performance. Many studies proposed several types of non-linear models using histopathological cell images (31-33), immunohistochemistry images (34-36), sonography and ultrasound images (37, 38), and different levels of gene expressions (39, 40).

Therefore, we designed the present study to determine the overall diagnostic accuracy of artificial intelligence systems that make a decision based on histopathology cell images for diagnosing breast cancer. Besides, the developed meta-analysis in-home tool which is available upon request, five tests were also developed to assess the publication bias within studies, namely Meta-MUMS DTA.

MATERIALS AND METHODS

Search strategy and study selection

In this study, we used (PICO) model (i.e., the patient problem (or population), intervention, comparison or control, and outcome) to define a well-focused question of research (1, 2). Where "P" refers to breast cancer disease, "I" refers to obtained histopathological cell images, "C" refers to several machine learning algorithms used to compare diagnosing outcomes with those of pathologists as the gold standard, and "O" refers to the performance results of the artificial intelligence systems including "true positive," "true negative," "false positive," and "false negative" assessment values.

To perform a rigorous systematic search that the most relevant evidence answering the research question could be extracted, the SCOPUS database was selected, and the following Boolean query was used up to December 15, 2018. The SCOPUS database demonstrated a comprehensive indexing sever that (i) covered 100% of MEDLINE coverage, 100% of EMBASE coverage, and 100% of Compendex coverage (3); (ii) included all of the engineering terms (e.g., individually, artificial intelligence systems), and (iii) was searchable with article types to omit the non-peer-reviewed studies (e.g., conference papers, editorials, letters, etc.). Only original or research articles were of interest in proposing the Boolean term. We excluded conference papers, letters and review papers as well as datasets or database of images.

Boolean query:

TITLE-ABS (image) AND (TITLE-ABS (rf) OR TITLE-ABS (random AND forest) OR TITLE-ABS (bayes) OR TITLE-ABS (support AND vector AND machine) OR TITLE-ABS (svm) OR TITLE-ABS (fuzzy) OR TITLE-ABS (machine AND learning) OR TITLE-ABS (artificial AND intelligence) OR TITLE-ABS (decision AND support AND system) OR TITLE-ABS (neural AND network)) AND (TITLE (breast AND cancer) OR TITLE (breast AND disease)) AND TITLE-ABS (accuracy) AND (TITLE-ABS (*pathol*) OR TITLE-ABS (histol*)) AND DOCTYPE (ar) AND NOT (TITLE (dataset)) AND NOT (TITLE (mammo*)).

For guaranteeing the quality of the systematic review and meta-analysis of diagnostic test accuracy (DTA) (41), the workflow of Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies (42) known as PRISMA-DTA statement, a PRISMA extension for DTA studies was followed.

Data extraction

The inclusion criteria for data extraction were set as, *(i)* availability of the original or research article, *(ii)* use of histopathological cell images, *(iii)* use of artificial intelligence systems for diagnosis, and *(iv)* provide enough data on evaluating the performance of the proposed predictive model (i.e., a set of values for accuracy, sensitivity, specificity or a set of values for "true positive," "true negative," "false positive," "false negative"). The ones that did not satisfy the abovementioned criteria were excluded from further analysis. Additionally, wherever needed, the values for accuracy, sensitivity, and specificity were converted using a Matlab script to generate the corresponding amounts for "true positive," "true negative," "false positive," "false negative."

The extracted information was screened for their relevance in terms of titles and abstracts by two independent researchers (MS and MS). Then the selected appropriate ones' full papers were retrieved, along with their data for further analysis.

Quality assessment

The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) was necessary for the evaluation of the quality of the included studies (43, 44). The quality assessment tool had almost 80% reliability power for at least ninety included studies. So, for the meta-analysis DTA studies that possess less than 90 investigations, the quality assessment of the researches would not be applicable.

.Statistical analysis

The importance of utilizing artificial intelligent systems based on histopathological cell images for diagnosing breast cancer has been evaluated by several measurement parameters incorporated in diagnostic test accuracy procedures. Among various software for DTA studies such as Meta-Disc (46) written in Microsoft Visual Basic 6, OpenMEE with Python/R architecture (47), and Open Meta-Analyst (48) with Python/R architecture, the last one was selected for assessment of the results. For this purpose, a new meta-analysis tool for DTA studies has been developed in Matlab 2013 environment (i.e., Meta-MUMS DTA tool). The results were checked with Open Meta-Analyst software (scripted in Python and R with only 64-bit support for Windows operating systems including Windows 7 and 8) (48) for consistency, and also the tool was improved for its user-friendliness and graphical user interface (GUI). The included statistic features were calculated using a 95 % confidence interval (CI) which included pooled diagnostic odds ratio (DOR) (with Cochran Q, inconsistency I^2, and τ^2, pooling of likelihood ratios (i.e., positive LR and negative LR), the pooled specificity (with χ^2 and inconsistency I^2), the pooled sensitivity (with χ^2 and inconsistency I^2), summary receiver operator characteristics (SROC), and area under the curve (AUC). In the existence of any types of heterogeneity (inconsistency I^2 >50%) in the random-effect model, the meta-regression and subgroup analyses would be performed if possible.

Moreover, the publication bias assessment feature was added to include Egger's regression-based test (49), Deeks' funnel plots (50), SVE (Smoothed Variance regression model based on Egger's test) (51), SVT (Smoothed Variance regression model based on Thompson's method) (51), and trim and fill combined with ln ω (logarithm of the diagnostic odds ratio) as suggested by Bürkner and Doebler (52). All statistical assessments were considered significant when the *p-value* was less than 0.05.

RESULTS

Database search

The searching retrieved 37 items. After screening the articles for their availability, 33 methods remained. Finally, three studies (53-55) were selected after going through the full texts for the existence of enough data relevant to histopathology cell images of breast cancer based on the inclusion criteria (Figure 1).

Characteristics and quality of the included studies

The studies' properties, their extracted data (i.e., TP, FP, TN, FN), and the used decision-making methodologies are tabulated in Table 1. As the study by Al Nahid et al. practiced six methodologies (as listed in Table 1) on their selected dataset, the number of included methods for analysis increased eight.

The quality assessment of papers was not carried out by QUADAS or QUADAS-2 tools (43, 44, 56, 57), as a recent research study did not recommend it for investigations less than 90 (45). In other words, a robust Monte Carlo simulation model has proved the reliability power of QUADAS-2 is almost 80% when an approximate number of included studies is 90. Hence, the measurement criteria for individual researches will remain unproven (45). However, one may find several articles demonstrating the quality of included studies using the abovementioned tool without taking in to account the threshold number (i.e., 90) for the researches (58, 59).

Figure 1. The PRISMA-DTA workflow for identification of included studies.

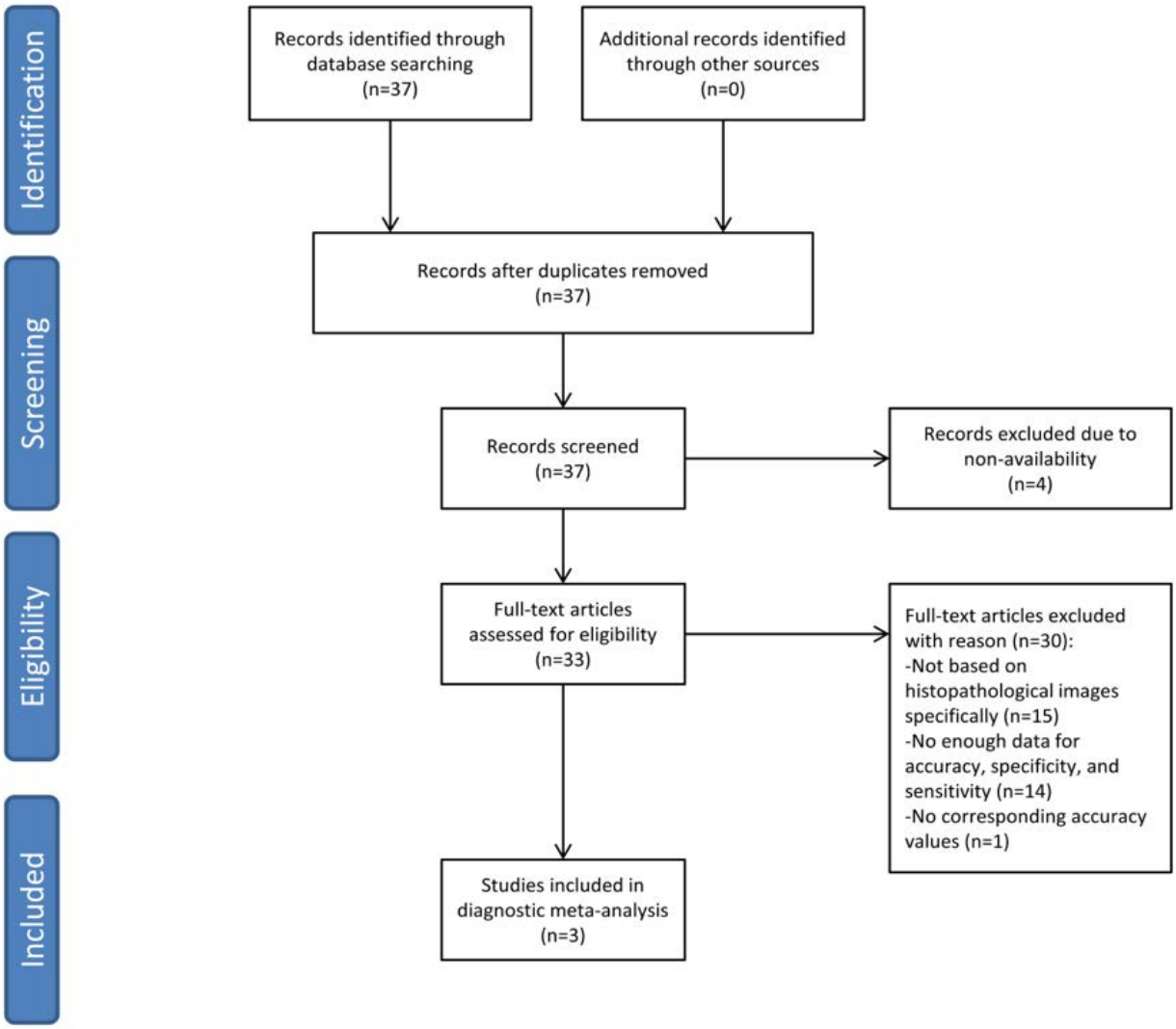


Table 1. Summary of the included studies with their characteristics.

Item	Author	Year	Country	Methodology	Total (cell)	TP	FP	FN	TN
1	Al Nahid A (55)	2018	Australia	MS-SVM	7909	5219	721	228	1741
2	Al Nahid A (55)	2018	Australia	MS-Softmax	7909	5367	429	282	1830
3	Al Nahid A (55)	2018	Australia	KM-SVM	7909	4995	820	208	1885
4	Al Nahid A (55)	2018	Australia	KM-Softmax	7909	5800	217	574	1318
5	Al Nahid A (55)	2018	Australia	OI-SVM	7909	5282	650	220	1757
6	Al Nahid A (55)	2018	Australia	OI-Softmax	7909	4623	659	290	2337
7	Wang P (53)	2016	China	SVM-CAGA	3600	1783	120	17	1680
8	Cruz-Roa A (54)	2018	United States of America	CNN	195	28	13	4	150

TP: True Positive, FP: False Positive, FN: False Negative, TN: True Negative, MS: Mean-Shift clustering algorithm, KM: K-Means clustering algorithm, OI: Original image without clustering, Softmax: a regression technique for decision layer, SVM: Support vector machine, CNN: Convolutional Neural Network, CAGA: chain-like agent genetic algorithm.

Meta-analysis for diagnostic accuracy

In the three studies including eight artificial intelligence-based approaches (i.e., MS-SVM, MS-Softmax, KM-SVM, KM-Softmax, OI-SVM, OI-Softmax, SVM-CAGA, and CNN) with a total of 11,704 cells the results for statistical analysis for diagnostic accuracy of breast cancer based on histopathology cell images are presented as sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odds ratio (DOR). Moreover, the area under the curve (AUC) of the summary receiver operator characteristic (SROC) is demonstrated to evaluate the total performance of each artificial intelligence-based diagnostic approach. The calculated values for pooled sensitivity, pooled specificity, PLR, and NLR were 0.95, 0.78, 5.25, and 0.06, respectively, as shown in figures 2 and 3. And, they were representative of the fact that breast cancer diagnosis based on artificial intelligence systems using the histopathological cell images performed great with the pooled F-score 0.86. Moreover, the overall performance of the approaches was assessed by the diagnostic odds ratios (the pooled DOR) and AUC with the values 88.15 and 0.9526, respectively (figures 4a and 4b). Due to the existence of significant and high heterogeneity of pooled DOR (i.e., I^2 = 95.6727) between studies, a simple method known as threshold effect analysis based on Moses-Littenburg regression (i.e., a fitted line using the difference of logit-transformed "true-positive" and "false-positive" rates versus their average (60, 61) or in other words the Spearman correlation coefficient of sensitivity and specificity (46, 62)) was carried out (Figure 4c). The results showed that the threshold effect could not the origin of the heterogeneity (Spearman correlation = 0.2857, *p value* = 0.4927). The meta-regression and subgroup analyses were not performed due to the lack of enough data.

Figure 2. The meta-analysis forest plots of (a) sensitivity and (b) specificity with 95% confidence interval.

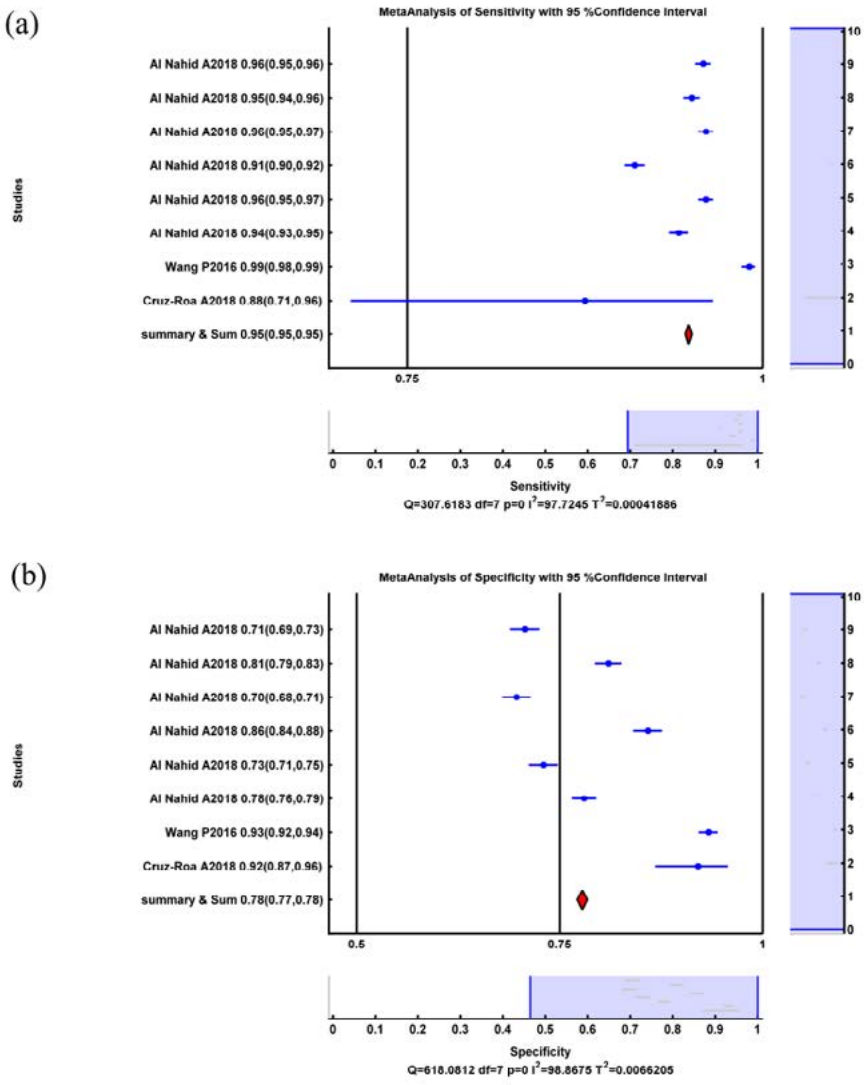


Figure 3. The forest plots of (a) positive and (b) negative likelihood ratios along with their statistical measurements (i.e., Q Cochrane, I^2, and T^2).

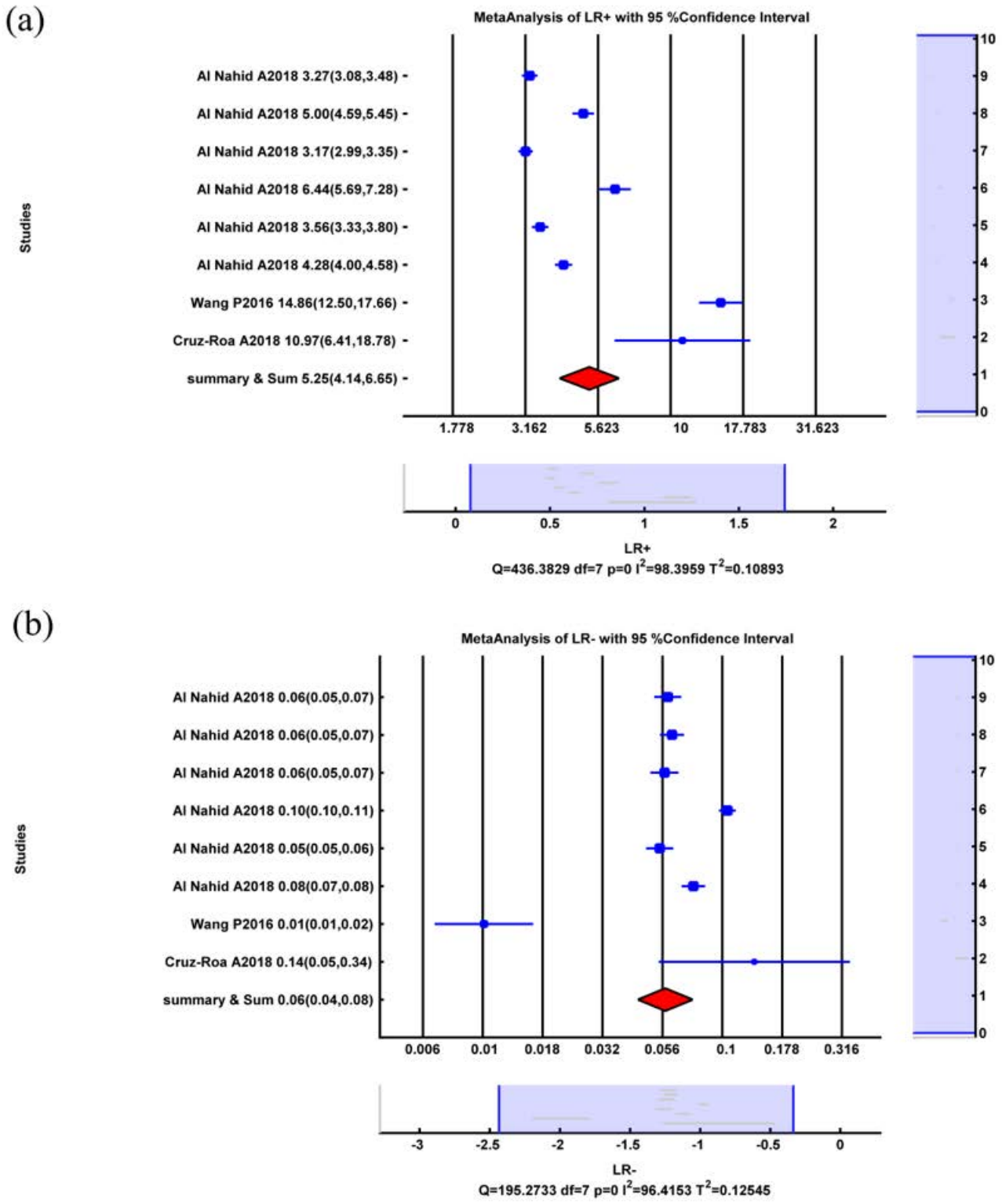
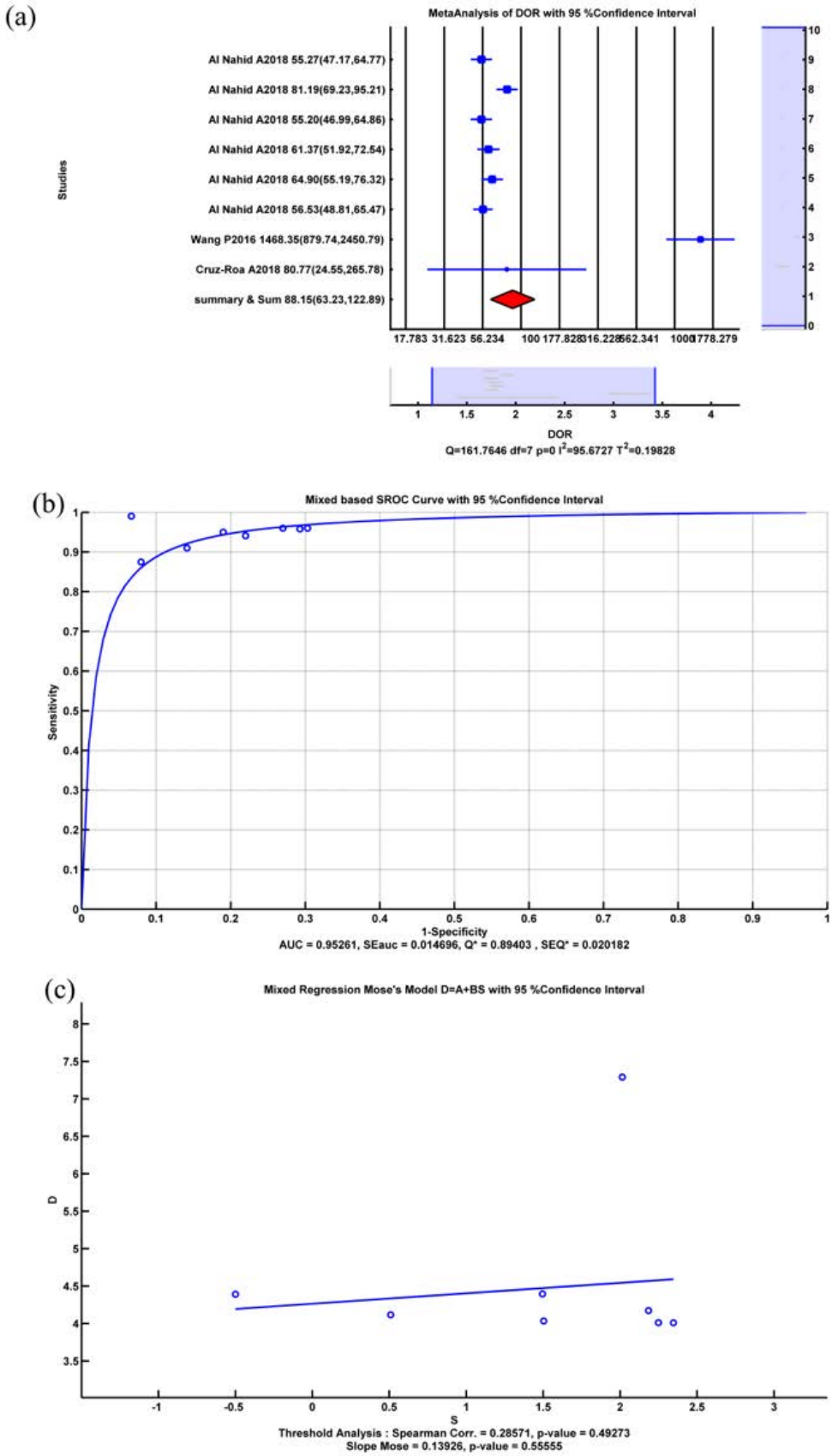


Figure 4. (a) Forest plot of diagnostic odds ratio of artificial intelligence systems in diagnosis of breast cancer (b) SROC of artificial intelligence systems in diagnosis of breast cancer, and (c) threshold analysis of eight artificial intelligence systems in diagnosis of breast cancer.



Publication bias

The publication bias of three included studies was performed by assessment test tools such as Egger's regression-based test, Deeks' funnel plot, SVE, SVT, and "trim-and-fill" methods combined with the logarithm of the diagnostic odds ratio. The results of the publication bias test are shown in Figure 5. The non-significant statistical *p-values* for Egger's (i.e., p value=0.51) and Deeks' (i.e., p-value=0.20) funnel plots were representatives of no potential publication bias.

However, the other remaining tests, including SVE (i.e., *p-value* = 0.02), SVT (i.e., *p-value* = 0.0003) with statistically significant *p-values* as well as for trim and fill methodology (showing three missing studies) demonstrated the existence of publication bias.

Figure 5. Publication bias test, (a) Egger's funnel plot (b) Deeks' funnel plot, (c) SVE funnel pot (d) SVT funnel plot, and (e) trim and fill assessment approach.

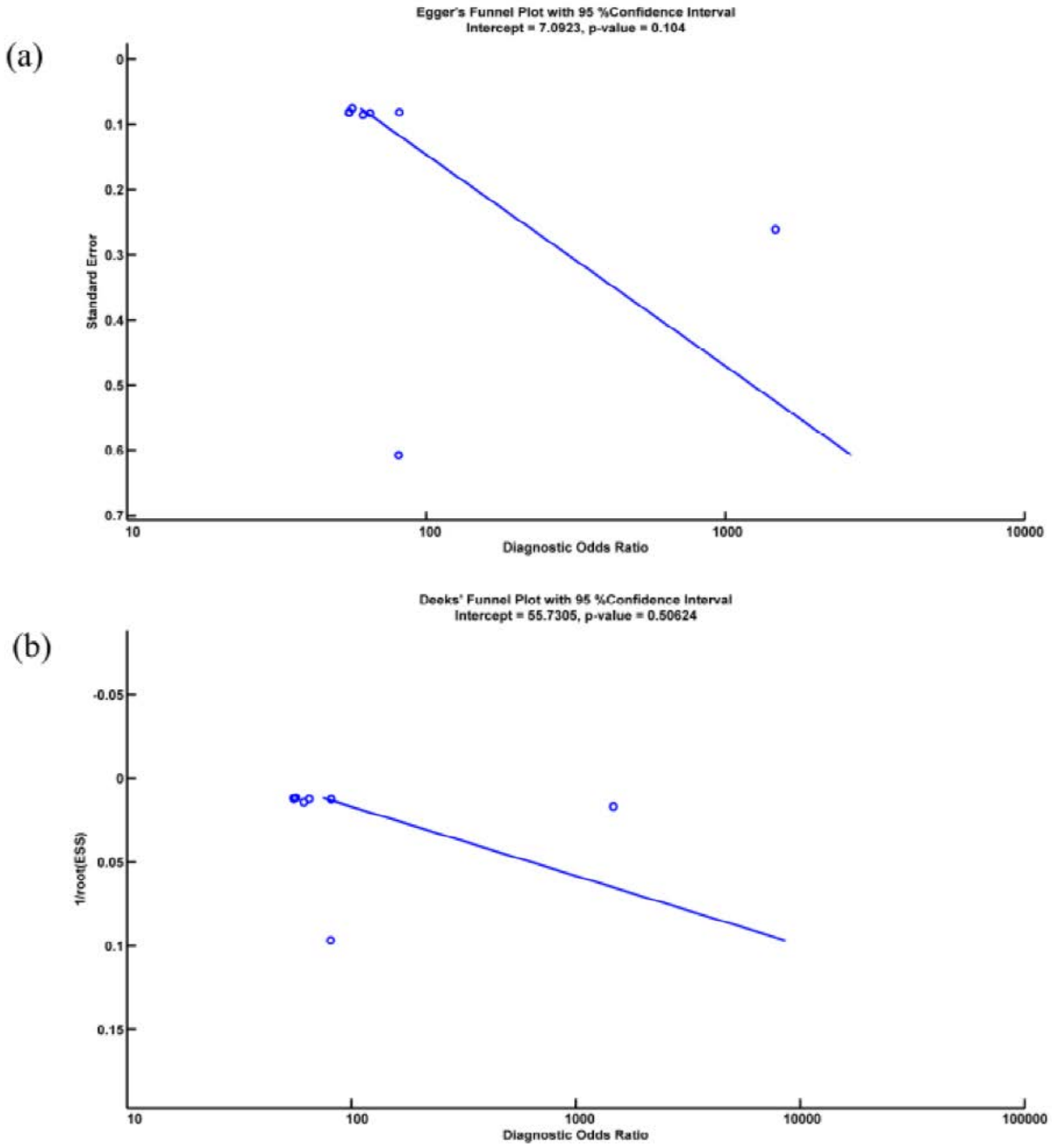
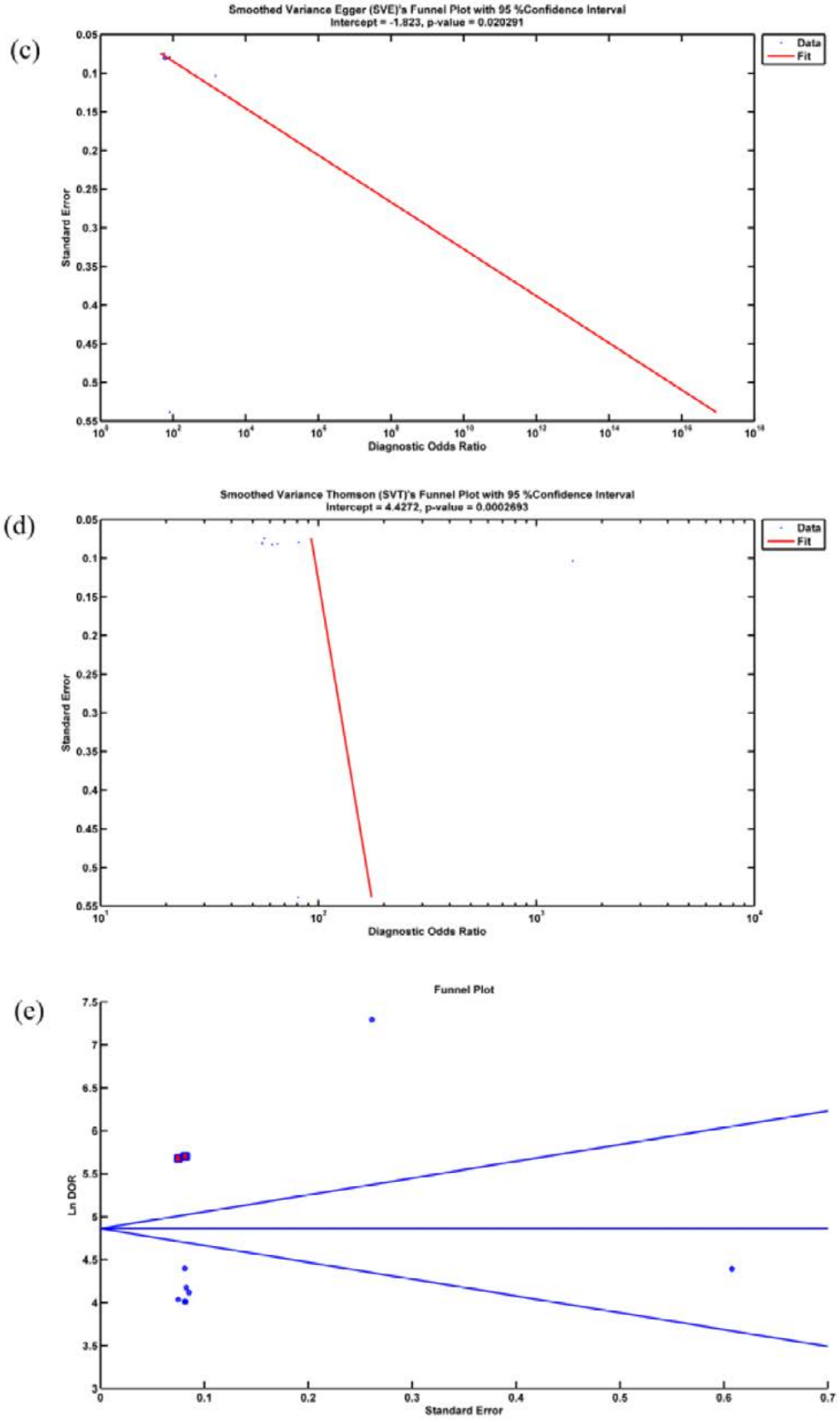


Figure 5. Publication bias test, (a) Egger's funnel plot (b) Deeks' funnel plot, (c) SVE funnel pot (d) SVT funnel plot, and (e) trim and fill assessment approach.



DISCUSSION

Breast cancer is the second most common cancer incidence in the world; however, the American Cancer Society has reported the mortality rate among women from 1989 to 2015 decreased by about 40% (63). Early detection of breast cancer using several methods remains however of literature's researchers' interest. Among the methodologies used for breast cancer diagnosis or prognosis, besides the pathologists' clinical decision making, taking advantage of artificial intelligence systems based on histopathological cell images seems crucial to avoid possible clinical misdiagnoses. And, the literature studies in this field only demonstrated their performance assessment results without including the results from previous studies. However, until now, no studies have been recommended to investigate the diagnosis value of artificial intelligence systems based on histopathological cell images.

The values of pooled sensitivity and specificity indicated that the overall rates of correct prediction of cancer image cells were higher in comparison to the total percentages of the right predictions of normal ones. And hence, the future directions of research are still open on proposing novel artificial intelligent systems mainly based on deep machine learning algorithms and robust, complete datasets of such images with uniform structures.

According to the formula $AUC =, \int_0^1 \frac{1}{1+\frac{1}{DOR\times(\frac{x}{1-x})}} dx$

The higher the value of the DOR (i.e., a single indicator of test performance) ranging from zero to infinity, the higher the value of the AUC, which ranges from zero to unity (64). The higher amount of DOR is indicative of the fact that the approaches can discriminate against the cancerous and healthy cell images with high overall accuracy; whereas, a DOR of 1.0 shows that the prediction methods do not have discrimination capability.

A PLR value of 5.25 implies that breast cancer patients have about a 5-fold higher chance for the positive status of cancer incidence in comparison to the normal ones. And the NLR value was 0.06, which showed the probability of breast cancer incidence is 6% and can be regarded as low to identify the cancerous cells from normal ones.

A statistically significant high level of heterogeneity existed among the studies' data, and it showed that it was not related to the threshold effect. Due to the lack of enough data, meta-regression and subgroup analyses were not applicable. Hence, the existing heterogeneity could be associated with the potential publication bias detected by SVE, SVT, and trim and fill approaches. Additionally, the trim and fill method revealed that three studies had been missed or not published in the literature due to lack of enough data, having negative results in comparison to others, or having exaggerating achievements in terms of reporting the highest performance of proposed models. Considering the results obtained by Bürkner and Doebler (52) using the simulation data, once again, we showed that trim and fill methodology was helpful with the current practical data in detecting the possible publication bias. However, Egger's and Deeks' funnel plots did not have that feature. Moreover, the SVE and SVT the improved modifications of Egger and Thomson were also able to identify the possible publication bias. Furthermore, SVE and SVT are suggested along with the trim and fill methodology for the assessment of publication bias.

The current systematic review and meta-analysis-DTA has several limitations as follows. First, Number of literature studies with adequate and complete statistical information sets, including TP, TN, FP, and FN as well as accuracy, sensitivity, and specificity, is not sufficient. And hence, the small number of studies was included for analysis. Second, the use of a uniform dataset will be advantageous for meta-analysis studies. The produced images may in fact derive from different devices with various settings and hence this may be pre-processed differently from study to study. Third, the lack of randomized clinical trials (RCTs) in the biomedical engineering field is another issue that needs much attention in future studies. Forth, the selection bias is an issue that may be occurred as there is no evaluation index for assessing the level of pathologists' professionalism. Fifth, the initial requirements for quality assessment of published articles showed that tools such as QUADAS and QUADAS 2 cannot be useful for less than ninety studies. Sixth, the lack of assessing tools for evaluating the two sections, including image processing and artificial intelligence systems that are vital issues that should be taken in to account for more details. Seventh, due to a lack of chance for publishing the articles with negative results, it will make any meta-analysis study prone to publication bias. And the ninth, no information was available for patients' demographic characteristics (i.e., gender, ethnicity, age) and a small number of publications during several years there cannot be provided no data entries for the country, continent, publication year, longitude, and latitude. Hence, meta-regression and subgroup analyses cannot be performed for defining the remaining sources of existing heterogeneity.

CONCLUSION

In this paper, a systematic review and meta-analysis diagnostic test accuracy evaluated the diagnostic value of artificial intelligence systems based on histopathological cell images in breast cancer. Although, results showed that the overall efficiency of a breast cancer diagnosis is high, further studies are essential to validate the outcomes of the

current study along with the large number of possible RCT studies including enough statistical and demographical information.

Conflict of interest

There is none to declare.

Acknowledgements

This study is part of a PhD thesis (No. 931507) and was approved and supported by Research Council of Mashhad University of Medical Sciences, Mashhad, Iran and carried out in Nuclear Medicine Research Center of Mashhad University of Medical Sciences, Mashhad, Iran.

References

1. Fineout-Overholt E, Johnston L. Teaching EBP: asking searchable, answerable clinical questions. Worldviews Evid Based Nurs. 2005;2(3):157-60.
2. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. BMC Med Inform Decis Mak. 2007;7:16.
3. Burnham JF. Scopus database: a review. Biomedical Digital Libraries. 2006;3(1):1.
4. World-Heath-Organization. Cancer 2018 [Access date: 17/04/2020]. Available from: http://www.who.int/news-room/fact-sheets/detail/cancer.
5. Rao JY, Apple SK, Hemstreet GP, Jin Y, Nieberg RK. Single cell multiple biomarker analysis in archival breast fine-needle aspiration specimens: quantitative fluorescence image analysis of DNA content, p53, and G-actin as breast cancer biomarkers. Cancer Epidemiol Biomarkers Prev. 1998;7(11):1027-33.
6. Truong K, Guilly MN, Gerbault-Seureau M, Malfoy B, Vielh P, Dutrillaux B. Evidence for in vitro selection during cell culturing of breast cancer: detection by flow and image cytometry. Cancer Genet Cytogenet. 1999;114(2):154-5.
7. Jitaree S, Phinyomark A, Boonyaphiphat P, Phukpattaranont P. Cell type classifiers for breast cancer microscopic images based on fractal dimension texture analysis of image color layers. Scanning. 2015;37(2):145-51.
8. Zaczek M, Kedra D. The use of computer-aided image analysis in the evaluation of epithelial nuclear area in the digestive tract of rats. Polish journal of pathology: official journal of the Polish Society of Pathologists. 1994;45(3):209-11.
9. Jaszcz-Gruchala A, Świątkiewicz I, Lackowska B, Gruchala A, Niezabitowski A. Comparative analysis of DNA content estimated by flow and image cytometry in breast tumor samples. Polish Journal of Pathology. 2001;52(1-2):15-9.
10. Olewniczak S, Chosia M, Kołodziej B, Kwas A, Kram A, Domagała W. Angiogenesis as determined by computerised image analysis and the risk of early relapse in women with invasive ductal breast carcinoma. Polish Journal of Pathology. 2003;54(1):53-9.
11. Kobos J, Dembski P, Taran K. The use of computer image analyses in the assessment of chosen histoclinical features of nephroblastoma cells in children. Polish Journal of Pathology. 2011;62(1):1-4.
12. Borkowski T, Dulewicz AM, Borkowski A, Piętka D, Radziszewski P. Clinical value of digital image analysis in the diagnosis of urinary bladder cancer, particularly in aggressive tumors: A preliminary report. Polish Journal of Pathology. 2016;67(2):122-9.
13. Jujjavarapu SE, Deshmukh S. Artificial Neural Network as a Classifier for the Identification of Hepatocellular Carcinoma Through Prognosticgene Signatures. Curr Genomics. 2018;19(6):483-90.
14. Takeuchi T, Hattori-Kato M, Okuno Y, Iwai S, Mikami K. Prediction of prostate cancer by deep learning with multilayer artificial neural network. Can Urol Assoc J. 2018.
15. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics Proteomics. 2018;15(1):41-51.
16. Zhou J, Li L, Wang L, Li X, Xing H, Cheng L. Establishment of a SVM classifier to predict recurrence of ovarian cancer. Mol Med Rep. 2018;18(4):3589-98.
17. Xiao LH, Chen PR, Gou ZP, Li YZ, Li M, Xiang LC, et al. Prostate cancer prediction using the random forest algorithm that takes into account transrectal ultrasound findings, age, and serum levels of prostate-specific antigen. Asian J Androl. 2017;19(5):586-90.
18. Ram M, Najafi A, Shakeri MT. Classification and Biomarker Genes Selection for Cancer Gene Expression Data Using Random Forest. Iran J Pathol. 2017;12(4):339-47.
19. Jain R, Mazumdar J. A genetic algorithm based nearest neighbor classification to breast cancer diagnosis. Australas Phys Eng Sci Med. 2003;26(1):6-11.
20. Alshamlan HM, Badr GH, Alohali YA. Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer classification. Comput Biol Chem. 2015;56:49-60.
21. Selvakumari Jeya IJ, Deepa SN. Lung Cancer Classification Employing Proposed Real Coded Genetic Algorithm Based Radial Basis Function Neural Network Classifier. Comput Math Methods Med. 2016;2016:7493535.
22. Hou Q, Bing ZT, Hu C, Li MY, Yang KH, Mo Z, et al. RankProd Combined with Genetic Algorithm Optimized Artificial Neural Network Establishes a Diagnostic and Prognostic Prediction Model that Revealed C1QTNF3 as a Biomarker for Prostate Cancer. EBioMedicine. 2018;32:234-44.
23. Bergholt MS, Zheng W, Lin K, Ho KY, Teh M, Yeoh KG, et al. In vivo diagnosis of gastric cancer using Raman endoscopy and ant colony optimization techniques. Int J Cancer. 2011;128(11):2673-80.
24. Shamsaee R, Fathy M, Masoudi-Nejad A. Extracting a cancer model by enhanced ant colony optimisation algorithm. Int J Data Min Bioinform. 2014;10(1):83-97.
25. Fallahzadeh O, Dehghani-Bidgoli Z, Assarian M. Raman spectral feature selection using ant colony optimization for breast cancer diagnosis. Lasers Med Sci. 2018;33(8):1799-806.
26. Chen KH, Wang KJ, Tsai ML, Wang KM, Adrian AM, Cheng WC, et al. Gene selection for cancer identification: a decision tree model empowered by particle swarm optimization algorithm. BMC Bioinformatics. 2014;15:49.
27. Sadoughi F, Ghaderzadeh M. A hybrid particle swarm and

neural network approach for detection of prostate cancer from benign hyperplasia of prostate. Stud Health Technol Inform. 2014;205:481-5.
28. Xi M, Sun J, Liu L, Fan F, Wu X. Cancer Feature Selection and Classification Using a Binary Quantum-Behaved Particle Swarm Optimization and Support Vector Machine. Comput Math Methods Med. 2016;2016:3572705.
29. Annavarapu CS, Dara S, Banka H. Cancer microarray data feature selection using multi-objective binary particle swarm optimization algorithm. EXCLI J. 2016;15:460-73.
30. Gao L, Ye M, Wu C. Cancer Classification Based on Support Vector Machine Optimized by Particle Swarm Optimization and Artificial Bee Colony. Molecules. 2017;22(12).
31. Sun D, Li A, Tang B, Wang M. Integrating genomic data and pathological images to effectively predict breast cancer clinical outcome. Comput Methods Programs Biomed. 2018;161:45-53.
32. Bernoux A, de Cremoux P, Laine-Bidron C, Martin EC, Asselain B, Magdelenat H. Estrogen receptor negative and progesterone receptor positive primary breast cancer: pathological characteristics and clinical outcome. Institut Curie Breast Cancer Study Group. Breast Cancer Res Treat. 1998;49(3):219-25.
33. Nikas JB, Low WC, Burgio PA. Prognosis of treatment response (pathological complete response) in breast cancer. Biomark Insights. 2012;7:59-70.
34. Dressler LG, Berry DA, Broadwater G, Cowan D, Cox K, Griffin S, et al. Comparison of HER2 status by fluorescence in situ hybridization and immunohistochemistry to predict benefit from dose escalation of adjuvant doxorubicin-based therapy in node-positive breast cancer patients. J Clin Oncol. 2005;23(19):4287-97.
35. Kim JW, Kim JH, Im SA, Kim YJ, Han HS, Kim JS, et al. HER2/CEP17 ratio and HER2 immunohistochemistry predict clinical outcome after first-line trastuzumab plus taxane chemotherapy in patients with HER2 fluorescence in situ hybridization-positive metastatic breast cancer. Cancer Chemother Pharmacol. 2013;72(1):109-15.
36. Guo Q, Chen K, Lin X, Su Y, Xu R, Dai Y, et al. A nomogram to predict HER2 status in breast cancer patients with HER2-borderline disease as determined via immunohistochemistry. Oncotarget. 2017;8(55):93492-501.
37. Keatmanee C, Makhanov SS, Kotani K, Lohitvisate W, Thongvigitmanee SS. Automatic initialization for active contour model in breast cancer detection utilizing conventional ultrasound and Color Doppler. Conf Proc IEEE Eng Med Biol Soc. 2017;2017:3248-51.
38. Giles EM, Parange N, Knight B. An Interprofessional Learning Workshop for Mammography and Sonography Students Focusing on Breast Cancer Care and Management Via Simulation: A Pilot Study. Acad Radiol. 2017;24(8):962-7.
39. Lin LZ, Cai MG, Dai YC, Zheng ZB, Jiang FF, Shi LL, et al. BRMS1 gene expression may be associated with clinico-pathological features of breast cancer. Biosci Rep. 2017;37(4).
40. Xiao B, Zhang W, Chen L, Hang J, Wang L, Zhang R, et al. Analysis of the miRNA-mRNA-lncRNA network in human estrogen receptor-positive and estrogen receptor-negative breast cancer based on TCGA data. Gene. 2018;658:28-35.
41. McGrath TA, Alabousi M, Skidmore B, Korevaar DA, Bossuyt PMM, Moher D, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. Syst Rev. 2017;6(1):194.
42. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, the P-DTAG, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018;319(4):388-96.
43. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
44. Schueler S, Schuetz GM, Dewey M. The revised QUADAS-2 tool. Ann Intern Med. 2012;156(4):323; author reply -4.
45. Venazzi A, Swardfager W, Lam B, Siqueira JO, Herrmann N, Cogo-Moreira H. Validity of the QUADAS-2 in Assessing Risk of Bias in Alzheimer's Disease Diagnostic Accuracy Studies. Front Psychiatry. 2018;9:221.
46. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol. 2006;6:31.
47. Wallace BC, Lajeunesse MJ, Dietz G, Dahabreh IJ, Trikalinos TA, Schmid CH, et al. OpenMEE: Intuitive, open-source software for meta-analysis in ecology and evolutionary biology. Methods in Ecology and Evolution. 2017;8(8):941-7.
48. Wallace BC, Dahabreh IJ, Trikalinos TA, Lau J, Trow P, Schmid CH. Closing the Gap between Methodologists and End-Users: R as a Computational Back-End. 2012. 2012;49(5):15.
49. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.
50. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005;58(9):882-93.
51. Jin Z-C, Wu C, Zhou X-H, He J. A modified regression method to test publication bias in meta-analyses with binary outcomes. BMC Medical Research Methodology. 2014;14(1):132.
52. Burkner PC, Doebler P. Testing for publication bias in diagnostic meta-analysis: a simulation study. Stat Med. 2014;33(18):3061-77.
53. Wang P, Hu X, Li Y, Liu Q, Zhu X. Automatic cell nuclei segmentation and classification of breast cancer histopathology images. Signal Processing. 2016;122:1-13.
54. Cruz-Roa A, Gilmore H, Basavanhally A, Feldman M, Ganesan S, Shih N, et al. High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection. PLoS One. 2018;13(5):e0196828.
55. Nahid AA, Mehrabi MA, Kong Y. Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering. Biomed Res Int. 2018;2018:2362108.
56. Schuetz GM, Tackmann R, Hamm B, Dewey M. Quality of diagnostic accuracy studies: QUADAS (Quality Assessment of Diagnostic Accuracy Studies Included in Systematic Reviews). Rofo. 2010;182(11):939-42.
57. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25.

58. Sims JI, Chau MT, Davies JR. Diagnostic accuracy of the Ottawa Knee Rule in adult acute knee injuries: a systematic review and meta-analysis. European Radiology. 2020.
59. McGrath TA, Frank RA, Schieda N, Blew B, Salameh J-P, Bossuyt PM, et al. Diagnostic accuracy of dual-energy computed tomography (DECT) to differentiate uric acid from non-uric acid calculi: systematic review and meta-analysis. European Radiology. 2020:1-11.
60. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Medical decision making : an international journal of the Society for Medical Decision Making. 1993;13(4):313-21.
61. Kardaun JW, Kardaun OJ. Comparative diagnostic performance of three radiological procedures for the detection of lumbar disk herniation. Methods of information in medicine. 1990;29(1):12-22.
62. Xu LH, Guo Y, Zhang XL, Chen JJ, Hu SY. Blood-Based Circulating MicroRNAs are Potential Diagnostic Biomarkers for Leukemia: Result from a Meta-Analysis. Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology. 2016;38(3):939-49.
63. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. CA: A Cancer Journal for Clinicians. 2017;67(6):439-48.
64. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol. 2003;56(11):1129-35.

*